Chemical reactions at the single molecule level

- Two molecules may have greater stability as a complex or as individual molecules.
- Transformation into an energetically less favorable state needs energy
- Some energy is required for the transition from one state to another even when the end state is the more favorable one
- For two molecules to undergo a chemical reaction, they must encounter each other in the correct orientation and they must have enough kinetic energy to overcome the transition energy barrier.
- In a gaseous mixture of two molecular species the average probability that a reaction takes place is

$$\chi = (c\ dt)\ X_1\ X_2$$

X_2: Nr. of molecules of species 2

$(c\ dt)$: Stochastic rate constant, depends on velocity, collision-path cross-section of molecules

Gillespie's algorithm

- Assume that M species can participate in N reactions.
- Need to know/estimate the stochastic rate constants for each reaction.
- Calculate the probability $P_{n,\delta t}$ that the next reaction will be reaction R_n, and it will occur δt time units from now.

- Idea: A reaction, once initiated, is instantaneous. Thus $P_{n,\delta t}$ is the product of the probability that no reactions occur in the period δt, and χ_n

$$P_{n,\delta t} = \chi_n e^{-\chi_0 \delta t}, \quad \chi_0 = \sum_{R=1}^{N} \chi_R$$

- Gillespie's algorithm picks the next reaction event by randomly sampling the $P_{n,\delta t}$ distribution for values of n and δt
- All reaction probabilities need to be recalculated after each event – improvements have been proposed

- Implementations in Mathematica, GillespieSSA (R package), Dizzy (used in our textbook)

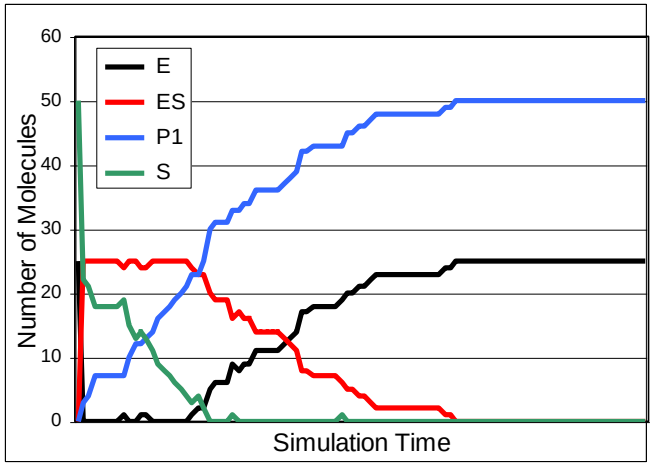
Example: stochastic Michaelis-Menten process

$E + S \rightarrow ES$ c1=1.0

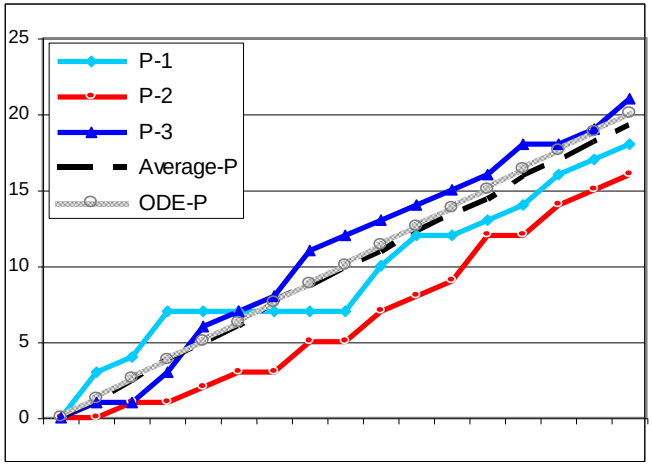
$ES \rightarrow E + S$ c2=0.1

$ES \rightarrow E + P$ c3=0.01

Initial condition: $E_0=25$, $S_0=50$, $P_0=0$, $ES_0=0$



A single simulation



The average of 50 simulations is very close to the continuous model's result

Recall: modeling of gene regulation

mRNA synthesis (transcription) – regulated by transcription factor(s)

Protein synthesis regulated by mRNA

Plentiful substrate for both transcription and translation

Transcription factors may also be regulated post-translationally

Modeling of transcriptional regulation may be broken into two steps:

- Determining promoter occupancy
- Writing the rate of mRNA synthesis as a function of promoter occupancy

Recall: continuous and deterministic description of average promoter occupancy

Single TF: $A + DNA_{free} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} A\bullet DNA$ → bound TF

$$\frac{d\,A\bullet DNA}{dt} = k_a A\, DNA_{free} - k_d .\, A\bullet DNA$$

equilibrium association constant

Steady state:

$$A\bullet DNA_{SS} = \frac{k_a}{k_d} A.DNA_{free} = K_A A.DNA_{free}$$

Promoter occupancy:

$$Y = \frac{A\bullet DNA}{A\bullet DNA + DNA_{free}} = \frac{K_A.A}{K_A.A+1}$$

Two cooperative TF:

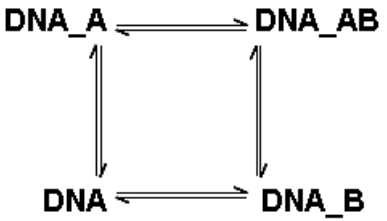
$$Y \approx \frac{K_A.A.K_B.B.K_q}{K_A.A.K_B.B.K_q+1}$$

cooperativity factor

Since there is only one DNA, continuous variables may not be justified.

A single-cell stochastic model of transcriptional regulation

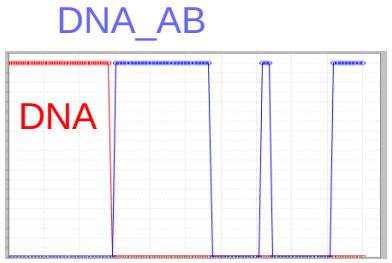
Two transcription factors cooperatively regulate the transcription of a gene



A_binds_DNA,	DNA + A	→ DNA_A,	k_{fA};
DNA_A_dissociates,	DNA_A	→ DNA + A,	k_{rA};
B_binds_DNA,	DNA + B	→ DNA_B,	k_{fB};
DNA_B_dissociates,	DNA_B	→ DNA + B,	k_{rB};
A_binds_DNA_B,	DNA_B + A	→ DNA_AB,	k_{fB2AB};
DNA_AB_releases A,	DNA_AB	→ DNA_B + A,	k_{rAB2B};
B_binds_DNA_A,	DNA_A + B	→ DNA_AB,	k_{fA2AB};
DNA_AB_releases B,	DNA_AB	→ DNA_A + B,	k_{rAB2A};

Differences from the continuous and deterministic description:

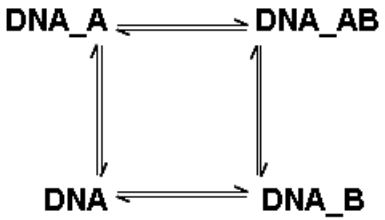
- The value of the variables can only be 0 or 1
- Instead of rates, we have reaction probabilities

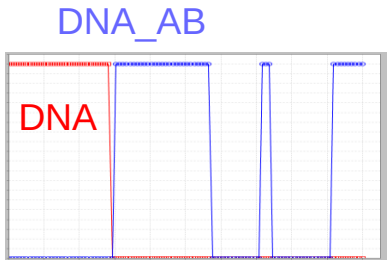


A single simulation

Probability of promoter occupancy

Two transcription factors cooperatively regulate the transcription of a gene



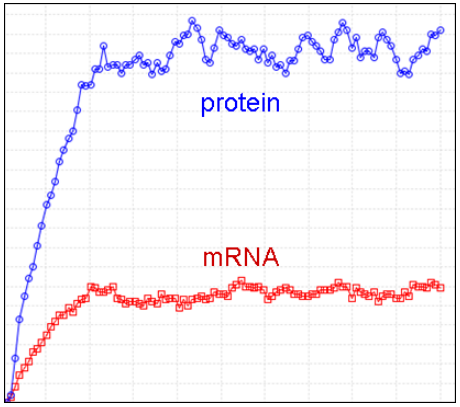


Result: the binary variable DNA_AB has a constant probability of being one in steady state.
One can calculate the probability of promoter occupancy as the ratio of the probability of DNA_AB to the sum of all probabilities.

The resulting occupancy is the same as the one coming from a continuous and deterministic description.

Next steps

- Recruitment of RNA polymerase
- Initiation of transcription
- Alternative splicing
- RNA editing
- RNA degradation
- Translation
- Protein degradation



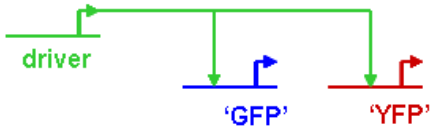
All processes modeled stochastically.

Ballpark parameter values known.

Can create a gene model that connects the protein copy number to the copy number of the two transcription factors.

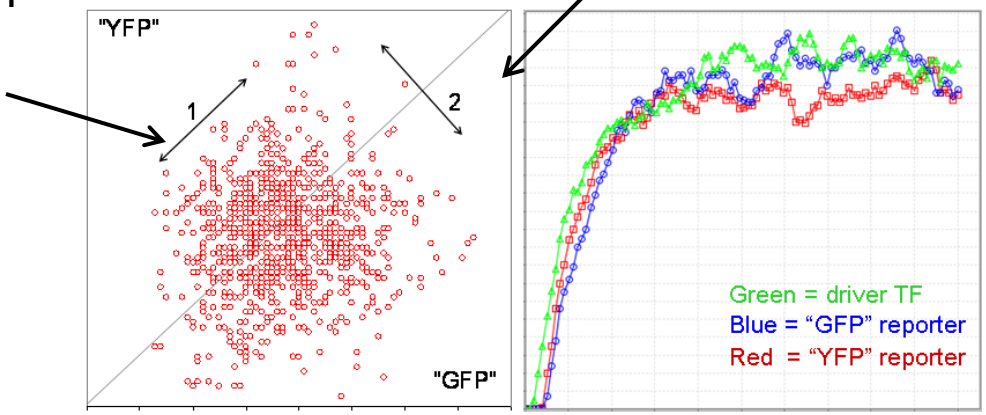
Find the expected amount of variability in gene expression levels in individual cells

- Synthetic circuit: a driver gene, two reporter genes regulated by the same promoter
- Assume that all three genes are regulated the same way, by cooperation of two (possibly identical) transcription factors.



Intrinsic transcriptional noise is quite large.

upstream noise



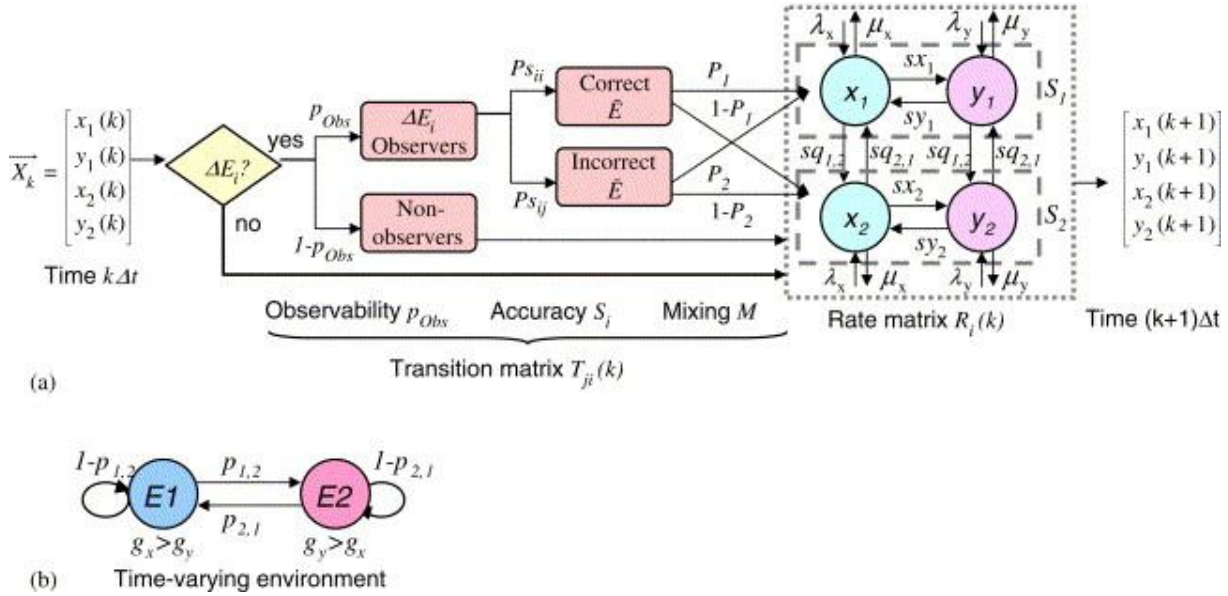
If the two reporter genes drive two competing processes, the outcome could be variable.

When is stochastic modeling necessary?

- When the variables have low integer values.
- When individual outcomes matter.
- E.g., when two regulatory proteins acting at low concentrations competitively control a switch point in a pathway, stochastic variations in their concentrations can produce probabilistic pathway selection, so that an initially homogeneous cell population partitions into distinct phenotypic subpopulations. This happens in many pathogenic bacteria.
- Conventional deterministic kinetics cannot be used to predict the statistics of regulatory systems that produce probabilistic outcomes. Rather, a stochastic kinetic analysis must be used.

Probabilistic strategies in microbial survival games

Environment fluctuates between two states, bacteria have imperfect environmental sensors, can switch between two states.



If the bacterial sensing is weak or has a long delay, random state changes are the best strategy (the Devil's compromise)
If environmental transitions are sensed but not the details of the new environment, bet hedging (mixing of strategies) is best.

D. Wolf, V Vazirani, A. Arkin, J. Theor Biol 2005.